Journal of Theoretical and Applied Information Technology
30th November 2019. Vol.97. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMPARATIVE ANALYSIS OF SUPERVISED AND UNSUPERVISED LEARNING ALGORITHMS FOR ONLINE USER CONTENT SUICIDAL IDEATION DETECTION

SERGAZY NARYNOV[1,2], DANIYAR MUKHTARKHANULY[1,2], ILMURAT KERIMOV[1], BATYRKHAN OMAROV[3-5]

[1]Alem Research, Almaty, Kazakhstan
[2]Suleyman Demirel University, Almaty, Kazakhstan
[3]Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan
[4]Al-Farabi Kazakh National University, Almaty Kazakhstan
[5]Kazakhstan Innovations Lab Supported by UNICEF, International Information Technology University, Almaty, Kazakhstan

Abstract

Suicide is one of the leading causes of death in most countries around the world; it is one of the three most common causes of death in a group of young people (15-24 years old), but so far no methods have been developed for diagnosing suicidal tendencies. In this connection, the problem of developing methods for identifying people prone to suicidal behavior is becoming especially topical. One of the directions of such research is the search for typological features of the speech related to suicide using the methods of mathematical linguistics, automatic text processing and machine learning. In foreign science, the texts of people that were motivated by suicide (mainly suicide notes) are studied using methods of automatic text processing (natural language processing), machine learning methods, and models that are constructed to allow to classify whether the text is related to suicide or not. It seems obvious that in order to develop methods for identifying people who are prone to suicide, it is necessary to analyze not only suicide notes (which are usually texts of small volume), but also other texts created by people who have committed suicide. The purpose of this work is to build a model of machine learning, apply teaching methods with and without a teacher, then select the most efficient algorithm for the task to classify whether the text is connected to suicide using comparative analysis.

Our research contributes to detection of depressive content that can cause suicide, and to help such people reach confident help from psychologists of national suicide preventing center in Kazakhstan. Obtaining highest result for 95% of f1-score for Random Forest (Supervised) with tf-idf vectorization model, in conclusion we may say that K-means (Unsupervised) using tf-idf shows impressive results, which is only 4% lower in f1-score and precision.

Keywords: *Random Forest, Sentiment Analysis, K-means, Machine Learning, Suicidal Ideation Detection.*

1. INTRODUCTION

According to data provided by the World Health Organization, more than 800,000 people die every year due to suicide, i.e. suicide is committed every 40 seconds, while, according to the available data, only 30% of those who committed suicide had previously reported their intentions [1]. Therefore, there is an objective need to develop methods aimed at identifying people that are prone to suicidal behavior and prevent their suicide attempt. The most valuable diagnostic tool that reveals the peculiarities of the personality psyche, including its tendency to suicidal behavior, is the analysis of its speech production, including the formal grammatical level, which is beyond the control of consciousness.

Social networks generate significant flows of information, which are characterized by a high level of dynamism and scale. The capacity of social networks is steadily increasing, and with it the speed of data transmission, and hence their processing, increases.

Based on the use of modern technologies for analyzing big data, the information obtained from social networks, which at first glance seems to be scattered, can be distributed according to a huge number of criteria – both common to certain groups of users and personally oriented.

This research compares supervised and unsupervised learning algorithms to identify suicidal ideation in popular social network (Vkontakte [2]) among adolescents and youth in Kazakhstan "VKontakte". VK (Originally VKontakte) is the second largest social network service in Europe after Facebook [3]. In the next section, we do literature review considering related works in this area. Section 3 describes the methodology of the research. Section 4 delves into experimental evaluation and analyzes the results in point of view of suicidal and non-suicidal post classification accuracy and choice of best algorithm for suicidal ideation detection.

The research motivation behind this study is to determine which type of machine learning algorithm is more appropriate and has high accuracy to detect users with suicidal tendencies on social networks. In their case, this gives researchers the opportunity to choose the optimal algorithm or model for applying machine learning techniques to a given task. In this paper, we will explore the answer of this research problem.

2. LITERATURE REVIEW

2.1. Pre-Suicidal Factors

The BBC Future review provides the latest and most extensive research into how social media, primarily Facebook, Twitter, and Instagram, affects our mental well-being [4].

Three billion people worldwide, or about 40% of the population, use social networks on the Internet [5]. We spend an average of two hours on them every day: we share posts, photos, and respond to friends' posts. Every minute, users of social networks send almost half a million tweets and photos to Snapchat.

Since social media has a significant impact on our lives, it is important to understand how they affect us.

STRESS. Social networks are a place where we often express our indignation about something, from poor-quality services to political problems. This allows us to let off steam, but turns our news feed into an endless stream of negativity.

Recent research tried to find out if social networks actually free us from negative emotions or, conversely, cause more stress [6-12].

ANXIETY. The researchers also tried to find out how social networks affect the overall level of anxiety. A study published in the journal Computers and Human Behavior (Computers and Human Behaviour) found that people who use seven or more networks have a three times higher general level of anxiety than users that use two social media platforms or less [13].

DEPRESSION. Although some past studies have found a link between depression and the use of social networks [14-16], new studies in this area suggests the opposite effect.

For instance, studies in which more than 700 students participated in showed that symptoms of depression, such as bad mood, feelings of inferiority, and despair, are associated with the quality of online communication [17].

For those to whom virtual communication brought mostly negative emotions, depressive symptoms were higher.

A similar study conducted in 2016 with the participation of 1,700 people demonstrated that the risk of depression and anxiety was three times higher among those who used several social media platforms [18-20]. Among the reasons cited by the researchers were, above all, virtual harassment and a distorted view of the lives of others [21-23].

However, scientists are also studying how social media can detect the symptoms of depression, which will help to promptly seek professional assistance [24-28].

By order of Microsoft, researchers surveyed 476 Twitter users and analyzed their social network profiles, paying attention to the style of messages, emotions, the type of interaction with other users, and signs of depressive behavior.

Using this data, they developed a questionnaire that allows us to predict the risk of depression in seven out of ten cases before the first symptoms appear.

LONELINESS. A study published by the American Journal of Preventive Medicine, where 7,000 people aged from 19 to 32 had participated in, showed that those who spend a lot of time on social networks feel twice as socially isolated. They lack the sense of belonging to a group, as well as interaction with others and full-fledged relationships.

According to researchers, for those people social networks crowd out personal relationships and make them feel lonely.

"An idealized view of the lives of friends and acquaintances can cause a feeling of envy and a false belief that they are happier and more successful. Such thoughts increase social isolation."

2.2. How Are Social Networks And Dysmorphophobia Related?

Social networks are regularly suspected of catalyzing bodily dysmorphophobia – dissatisfaction with one's own body and even hatred towards it, associated with the search for contrived deficiencies, eating disorders and other mental difficulties. The essence of the claims comes down to the fact that the photos on Instagram create a distorted body image, so that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

one's own one gradually begins to seem “somehow different” to a person [29-31].

Plastic surgeons are increasingly saying that social networks are provoking more operations. Today, patients come to them not with photos of celebrities, as before, but with snapchat filters, which makes it possible to speak of a new type of dysmorphophobia. On the other hand, body positive does its job: social networks become a platform through which people whose appearance was not represented in the media space before, tell the whole world about themselves – and then appear in advertising campaigns, go to the podium and by their own example prove that body worthy of respect.

Scientists have determined that those who showed signs of depression, behaved in social networks as follows:

Symptom 1

They use social media to compare themselves with “better” version of other users. Most often, this trend can be traced on Instagram, when users post only successful photos and demonstrate a successful life, leaving reality behind the scenes;

Symptom 2

They regularly use social networks and this comes to the level of “addiction” (this condition was determined by the results of a survey in which students answered “yes” to the following statements: “you tried to minimize the use of social networks”, and “this had a negative impact on your work/study”).

Symptom 3

They felt worried after they were noted on the failed photos on social networks. Often this caused students anxiety and even aggression.

Symptom 4

These students rarely publish photos with other people. The reason why people with depression are less likely to publish pictures with others, according to the authors of the study, is that individuals with this psychological problem often tend to isolate themselves from others.

2.3. Applying AI in suicidal ideation detection and suicide prevention

What is most disturbing is the fact that suicides are a common cause of death among adolescents and young people under the age of 25 [32-37].

Attempts to teach technology to fight suicide have been going on for quite some time: Back in 2007, a group of researchers attempted (rather successfully) to analyze the records of MySpace.com users in order to identify those who are on the verge of suicide. Among other things, it is known that machine learning algorithms are much more efficient than professional doctors in distinguishing a suicide note from a fake one(78% accuracy against 63%), and suicidal tendencies can be indicated, for example, by the duration of vowels during a conversation. The search for these does not arise from scratch: as a recent study showed, over fifty years of studying suicide, scientists have not made any visible progress in predicting it [38-39]. Traditional risk factors identified over the past half century - depression, stress, use of psychoactive substances - provide accuracy not higher than random guessing.

The main problem of previous works is considered to be an extremely narrow methodology: most of them concern only one risk factor and do not take into account the complexity of their combinations at all.

Propensity to suicidal behavior can be determined by using automatic analysis of brain activity. It was succeeded by the American scientists who had applied methods of machine learning for the analysis of results of fMRI scanning of potential suicides [40]. The new algorithm identified suicidal people with an accuracy of 91 percent. The article was published in the journal Nature Human Behaviour [41]. Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth.

Recent work indicates that by analyzing patterns of brain activity, we can identify the neurobiological prerequisites for determining the concepts of the real world — how concrete concepts are formed and stored in our mind. For example, the understanding of the concept of “home” is associated with the activity of the parahippocampal gyrus, which is responsible for processing information about scenes of reality. The authors of the new work decided to check whether there are differences in brain activity in defining various emotional concepts between healthy people and people with suicidal tendencies - and, if they exist, is it possible to determine a person's suicidal tendencies [42].

Applying machine learning based on the use of the naive Bayes classifier, scientists identified voxels (three-dimensional image units) of the differences in activation when comparing 17 healthy people and 17 people with suicidal thoughts. The researchers managed to identify six concepts - "death", "cruelty", "problem", "carelessness", "good" and "praise" - where activation of them is the most different between the two groups.

Indiana University conducted a longitudinal study of 51 women with diseases such as bipolar disorder, depression, schizoaffective disorder, and schizophrenia. In all patients during treatment, the standard index of suicidal thinking (SI) was determined. They also determined the level of expression of various genes, choosing those for further study, the expression of which is most different in women without suicidal thoughts and participants with a maximum SI.

Using probabilistic analysis of available genetic information bases in selected genes, the most reliable biomarkers of suicidal susceptibility were identified. The effectiveness of the proposed diagnostic methods was estimated by the area under the ROC curve (ROC AUC). For UP-Suicide, this figure was 82 percent when predicting SI and 78 percent when predicting future hospitalizations for suicidality. [43]

An important topic of the June (2017) issue of the British Journal of Psychiatry is the prediction and prevention of suicides. Carter and colleagues (Carter et al.) Performed a systematic review and meta-analysis of 39 risk scales used in predicting suicidal behavior. As a result of the study, the positive predictive value of these scales (ranging from 5.5% for suicide to 35.9% for a combined outcome) was considered too low to be used to make decisions about clinical interventions. Alternatively, the authors propose a clinical assessment in order to identify variable risk factors and apply specific interventions intended for individual subgroups of self-harming individuals [44].

The breakthrough project, carried out under the guidance of scientists from the University of Florida, has dramatically increased the ability to predict suicides - it allows you to predict with 80% accuracy which patient will make a suicide attempt in the next two years. A new article by the main artist Jessica Ribeiro was published in Clinical Psychological Science. According to the results of the study, machine learning can, with 80–90% accuracy, predict the probability of a suicidal attempt for a particular person two years in advance. As the suicidal attempt approaches, the accuracy of the algorithms increases even more, for example, a week before the attempt, the accuracy rises to 92% (for patients of the general hospital) [45].

Ribeiro’s research is particularly impressive compared to a recently published review of 50 years of research in the field of suicide prediction (by Joseph Franklin), which showed no real progress in this area. The Ribeiro project was born out of Franklin's shocking result. Ribeiro, along with Franklin, and Colin Walsh processed a large amount of data (about 2 million anonymous electronic patient records in Tennessee). In this database, more than 3200 people who committed a suicide attempt were identified. This was key information: similar medical records of thousands of suicidal people.

After studying all these data, machine learning algorithms could “find out” which combinations of factors in the stories were predicted by future suicidal attempts with the greatest accuracy. “The machine finds the optimal combination of risk factors,” says Ribeiro. “The main thing is how this algorithm and these variables interact with each other as a whole. This kind of work allows us to apply algorithms with hundreds of data points from a medical history and potentially reduce them to clinically relevant information. ” Such meaningful information can be used to develop an “alarm system” for clinicians who identify patients at risk for suicidal behavior. According to research, about 60-90% of people who died as a result of suicide, visited the doctor that year, but the clinicians did not see the approach of suicide. Currently, work is underway in the United States to create national and international e-case histories for subsequent analysis using a machine learning algorithm to identify individuals at risk of suicide. The US Department of State and the Department of Veterans Affairs have already launched this program for their own database. [45.]

The Korean government decided to fight the growing number of suicides in the country using a Facebook program called AI Saving Lives. It uses machine learning algorithms to track and analyze user activity in social networks and instant messengers. The Facebook algorithm is also able to detect video recognition, which suicides usually make before they die. When such content is detected, the program sends a warning to the user and his friends [46].

Scientists from the University of Cincinnati (University of Cincinnati), University of Colorado at Denver (University of Colorado Denver), University of Southern California (University of Southern California) and Princeton University (Princeton University) have developed an algorithm that can determine whether a person is prone to suicide. The research results are published in the journal Suicide and Life-Threatening Behavior [47].

The machine learning algorithm was able to distinguish patients with suicidal tendencies from mentally ill and healthy people with an accuracy

Journal of Theoretical and Applied Information Technology
30th November 2019. Vol.97. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of up to 93%. "These results convincingly prove that advanced technologies can serve as a decision support tool that will help clinicians and social workers to recognize and prevent suicidal behavior," says one of the authors of the study, Professor John Pestian. "You see tremendous support from technology in medical institutions, but those who work with mentally ill people do not receive it in such volumes. [47].

Scientists have long taught computers to diagnose suicidal tendencies. Eleven years ago, researchers created software that recognized emotions in suicide notes, eight years ago, machines distinguished real notes from simulations. In 2015, scientists analyzed people no longer posthumously, but during a doctor's appointment, but on the example of a small sample — 60 people — and in one medical center. New article is devoted to the first distributed multicenter study.

It was attended by 379 people from the Medical Center of the Children's Hospital of Cincinnati, the Medical Center of the University of Cincinnati and the Princeton Public Hospital. Participants belonged to one of three groups: suicidal, mentally ill but not suicidal, and healthy people (control group). Those who were within 24 hours or went to an ambulance or a psychiatric hospital in connection with an attempt or intention to commit suicide, were mentally ill - people who received an appropriate diagnosis were considered prone to suicide.

Patients underwent standardized tests for depression, the severity of suicidal manifestations, and the severity of mania. Then - an interview in which the doctor asked the questions "Do you have hope?", "Do you have secrets?", "Do you have fears?" And others. The scientists recorded video conversations. They deciphered the interviews, created a "dictionary" of key words and sound characteristics, and on a part of this data they trained a computer learning algorithm, the so-called "support vector machine", the purpose of which was to find a hypothesis with the smallest real error.

Then the computer was "fed" the remaining recordings and interview transcripts. It turned out that the machine distinguishes patients from three groups with an accuracy of at least 70%. The algorithm could analyze both one linguistic or acoustic component of speech, and both together. When comparing suicidal people with a control group, the algorithm reached an accuracy of 93 (text only), 79 (audio only) and 92 (text + audio) percent. He distinguished suicidal from mentally ill people in 79 (text), 76 (audio) and 81 (text + audio) percent of cases. Interestingly, the use of acoustic information in some situations increased, and in others reduced accuracy.

Studies of recent years with the use of these software tools analyze texts of diaries, letters (including those published on the Internet), and lifetime interviews of people who have committed suicide [48]; an overview of this kind of work is presented in D. Lester [49].

It should be noted that these works do not pose the task of creating methods for diagnosing a tendency to suicidal behavior based on a quantitative analysis of speech products, but only states that there are statistically significant differences in the texts of suicides and non-suicidal persons. The main text analysis tool used in these works is the LIWC program. The work of Mulholland M., Quinn J. [50], performed on the material of the lyrics, also poses the task of developing a mathematical model that classifies the texts as belonging to suicides or to individuals in the control group. However, the study significantly expanded the list of text parameters (TTR, the proportion of words in some parts of speech, the proportion of verbs in the passive voice, the proportion of words in certain semantic classes, the grammar of words). Texts were tagged using modern means of automatic language processing. Using the methods of machine learning, a classifier was built, the accuracy of which was 70.6%. Of course, the accuracy of the classifier is far from desirable, but the results show the fundamental possibility of solving the problem of diagnosing the risk of suicidal behavior based on the analysis of quantitative analysis of texts using NLP methods and mathematical statistics. As the authors of the study rightly point out that in order to improve the accuracy of the model, it is necessary to expand the body of texts and the set of parameters for analysis.

3. SELECTION OF FEATURES OF A SOCIAL NETWORK USER FOR BUILDING A MACHINE LEARNING MODEL

Undoubtedly, social networks have become an integral part of the lives of billions of people and it is difficult to imagine life without them. In addition to communicating with friends, family, acquaintances, social networks are used to read current news and are a huge platform for sharing information. It is not a secret that both small entrepreneurs and large corporations operate on social media. With the active development of the Internet, social networks have become one of the key marketing tools. Indeed, among the users of social networks there is an audience of all ages and different segments of the population.

On social networking sites, you can successfully promote your product without deep analysis of the audience, but you can get much more benefit if the target user has demographic attributes, interests, hobbies, or certain behavior, based on social network data. Such information allows, for example, to offer personalized product offers that are more suitable for this user, or to identify potential buyers and move on to more active sales methods.

As a rule, in many services, each user has a name, age, location and a short description of interests, or links to other social networks. But this is sometimes not enough or the information may not be correct. All this requires improvements in data collection and processing systems.

Basic user classification model

Data that represents information about the user on social networks can be divided into 4 main types: profile, behavior, text content (messages, publications, comments) and data from the social networks themselves (for example, time of visits, number of subscribers \ subscriptions).

These our types are enough to extract features of a general-purpose classification model. We face the challenges:

- General assessment of the relative impact of the trait, reliability and generalization of potential opportunities for user classification.
- Exploring the value of linguistic information for classifying users.

Signs from user profile

Many services, by default, display profile information, such as user name, location, and a short summary about the user.

In addition, all popular social networking APIs provide access to other information, such as the number of subscribers, friends, publications, etc.

However, during the data collection from social networks you can face a number of problems:

Restricted access or blocking for automatic collection of information.

Data privacy – often users set privacy settings, which makes many profile attributes inaccessible from the outside.

The weak structure of the data - some social networks either do not have an API, or with large restrictions, which makes data collection very problematic. To do this, you must “manually” get HTML and parse the page structure.

Behavior

Characteristics of user behavior include a number of statistics of interactions with the service: the average number of messages per day, the number of responses, and the number of repeated publications (reposts), etc. Such data may be quite suitable for training the model. For example, statistics show that people who rarely make publications have more unique content, comparable to those who do it more often and publications contain a URL to third-party sites.

Text content

Text content has information about the main topics of interest to users. Simple text information helps to classify users by comments, blogs, conversations and search sessions.

Next, we analyze the basic properties of textual content.

Modeling by topic

Taking into account n classes, where each ci represents a set of users Si. Each word assigned to at least one user is assigned a score for each of the classes. The value of the score is calculated by the conditional probability of the class according to the formula:

$$proto(w, c_i) = \frac{|w, S_i|}{\sum_{j=1}|w, S_i|}$$

where | w, Si | this is the number of words issued by w for all users of the class ci. For each class of users, k thematic words with a high score are saved. The n ∗ k thematic words are collected from all classes and serve as signs representing a specific user: for each thematic word wp, an estimate is assigned to user u.

$$f_{proto(w,c_i)} = \frac{\sum_{wp\ c} WP|u, wp|}{\sum_{wpW_u}|u, w|}$$

Wp is a set of thematic words for class c.

Hashtag classification

In many social networks you can use hashtags to indicate the main topic of the publication. Most often the same or similar hashtags are used to simplify information retrieval among similar publications. Classification by hashtags occurs in the same way as for subject phrases. Given that the set is Si for class ci, where it will contain all the hashtags from the publication. Then you can get a set of themes of hashtags by applying Formula 1. Ultimately, you need to save hashtags with high marks for each class, and calculate the attributes by formulas 2 and 3.

Interactions

Relationships established by the user with other users of the social network, his answers, reposts, and publications he likes have a large role in classification.

Friends

It is easy to see that people who are fond of cars are more likely to add like-minded friends to their friends and join car-related groups. It can be assumed that users of other classes can also share specific accounts of friends.

In order to load a set of classes of profiles of friends F, you can use the same classification mechanism for subject phrases, namely formula 1. Then you need to get the following characteristics for this user u: the number of profiles in [FF that the user is subscribed to; the percentage of F profiles that are subscribed to u; The percentage of accounts subscribed to u.

Answers, comments and reposts

Here, the same principle as with friends, the sign shows that the user has a tendency to comment, make reposts, add to the selected content of people or groups of interest. It is possible to classify by the same principle, by thematic words, and hashtags.

Communities

In the real world, it is quite typical for humanity to unite in various communities of interest, occupation, social circle. The same picture is observed in social networks.

An analysis of the social network user communities is also an important classification criterion. Information about communities, the social network at the global level finds application in the systems of recommendations, filtering and spam and many other applications. As a rule, communities in social networks are internally categorized, which greatly simplifies the task of classifying users who join them.

4. METHODOLOGY

Before classifying information as suicidal or depressive, it is necessary to define the criteria of "danger". One of the solutions is the definition of a set of keywords. It is a method of determining the types of information and is applied in the developed software package. For the definition a set of keywords was compiled, which was used to analyze information on the social network VKontakte. The software package based on the presence or absence of the specified keywords in the text concludes that the text is suitable for further research.

The implementation of obtaining information may vary depending on the source of information, but maintain the general principle of its construction. The main purpose of the part of the software responsible for obtaining information from open sources is to perform actions quickly and efficiently. To achieve maximum performance, you must use the built-in methods for obtaining information from sources (API), if any. If there are no such methods, then it is necessary to obtain and extract the necessary information from HTTP requests.

There are three separate modules of the software package:

1. Information collection module - is responsible for receiving information from open sources and transmitting it for further processing;
2. Keyword search module - is responsible for finding keywords in a large amount of information;
3. Document ranking module - is responsible for determining whether the information is dangerous.

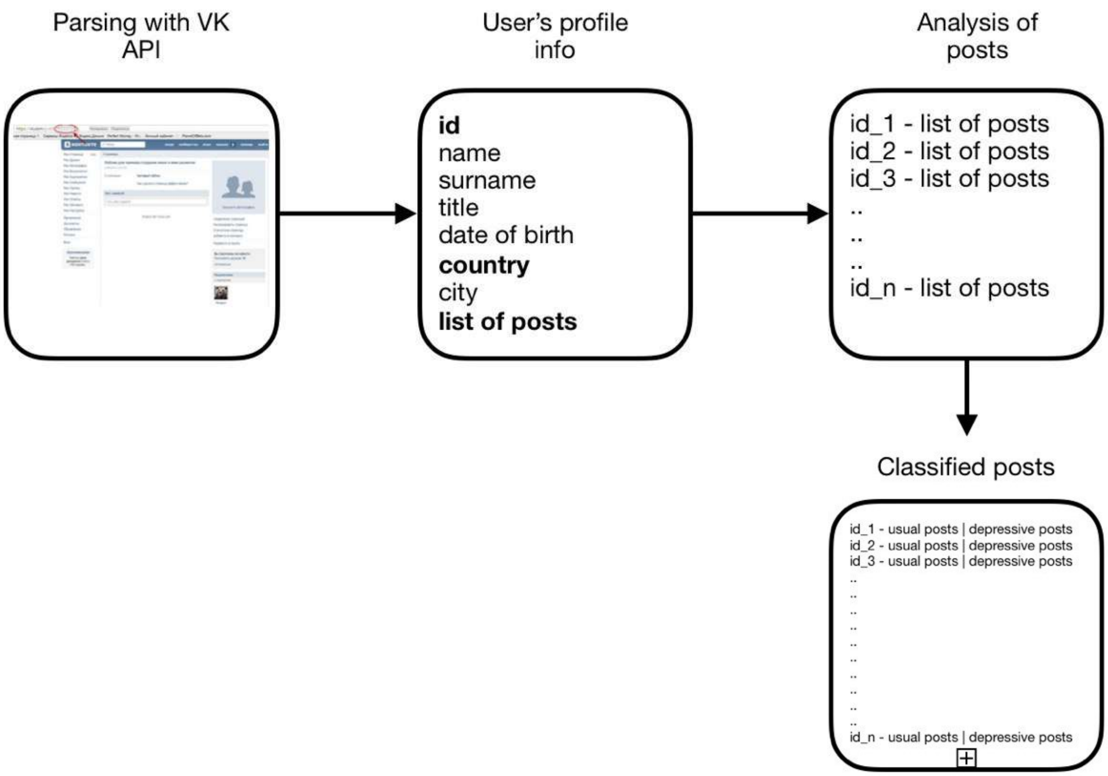

Figure 1. Data Division Of Training And Testing Dataset

We collected 35,000 messages in Russian for people diagnosed with depression at various levels: severe, chronic, manic, persistent depressive disorder, and so on. We also used about 50,000 personal posts from social networks with negative attitudes on various topics to create a homogeneous text base, divided into a training base for learning algorithms and test bases for evaluating the performance of algorithms. Training and testing dataset divided by 74.5% to 24.5% as illustrated in Figure 2.

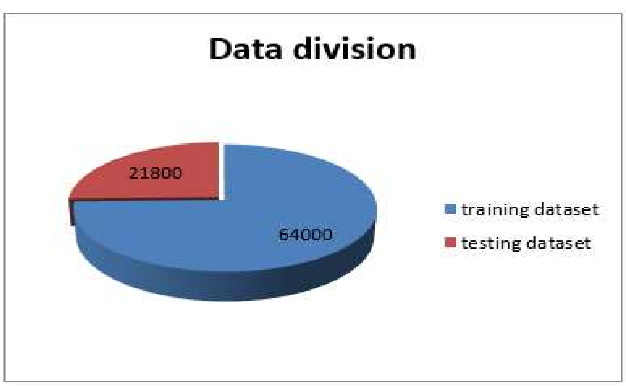

Figure 2. Data Division Of Training And Testing Dataset

Based on a thorough analysis of the text corpus and the main themes and emotions, our psychologists identified three possible causes of suicide:

1. The need for support is a desperate attempt to draw the attention of other people to their mental state.

2. Avoidance - the inability to tolerate any further intolerable heartache, guilt, or shame for socially unacceptable acts.

3. The protest is a protest against difficult family problems, often associated with the expression of emotions of anger and accusation. A written note is often addressed to a specific person or group of people.

Suicide is not an impulsive act, and preparation for suicide can last for about a year, during which a person will show signs of his condition. We use the analysis of sentiment to detect this dangerous period.

5. EXPERIMENTAL EVALUATION AND ANALYSIS THE RESULTS

We used Random Forest, Gradient Boosting

Journal of Theoretical and Applied Information Technology
30th November 2019. Vol.97. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[4] to detect depressed publications with supervised learning. While for teaching without a teacher, we used K-means with a 2-cluster model.

5.1. Data Preprocessing

First, all texts were lemmatized — the process of deleting only endings and returning the base or vocabulary form of a word, which is known as a

Accuracy 0.895081766399216

Metrics	precision	recall	f1-score	support
0	0.96	0.90	0.93	12212
1	0.75	0.89	0.81	4115
avg / total	0.91	0.90	0.90	16327

lemma. For the lemmatization of words in the context of the Russian language, the lemmatizer "MyStem" from Yandex was used, since it demonstrated excellent results. Subsequently, the nltk library for stop words was used to remove the stop word, hence reducing potential noise in the data. Numbers, special characters, not Cyrillic letters have also been deleted.

Secondly, the pre-processed texts were vectorized — the process of representing texts in a vector space for arithmetic operations on the entire data structure. Vector view saves time. For vectorization of texts, the TF-IDF and Word2Vec models were used.

TF-IDF stands for Term Frequency-Inverse Document Frequency, which basically indicates the importance of a word in a package or data set. TF-IDF contains two concepts: term frequency (TF) and reverse document frequency (IDF)

Word2vec is a deep learning technique with a two-layer neural network. Google Word2vec takes data from big data and converts it into vector space. Word2vec basically puts a word into feature space in such a way that their location is determined by their meaning, that is, words that have a similar meaning are grouped together, and the distance between two words also has the same meaning.

5.2. Supervised Learning Algorithms

For the experiment 2 best algorithms were tested according to our previously published work, as well as with the vectorization of tf-idf [4]:

1. Gradient Boosting with word2vec
2. Random forest with word2vec
3. Gradient Boosting with tf-idf
4. Random forest with tf-idf

5.2.1. Gradient boosting with word2vec

The result of applying Gradient Boosting algorithm as following:

Accuracy 0.9056777117657867

Metrics	precision	recall	f1-score	support
0.0	0.95	0.92	0.93	11811
1.0	0.80	0.87	0.84	4516
avg / total	0.91	0.91	0.91	16327

Precision: 91%
Recall: 91%
F1-score: 91%

5.2.2. Random forest with word2vec

The result of applying Random Forest with word2vec algorithm as following:

Precision: 91%
Recall: 90 %
F1-score: 90%

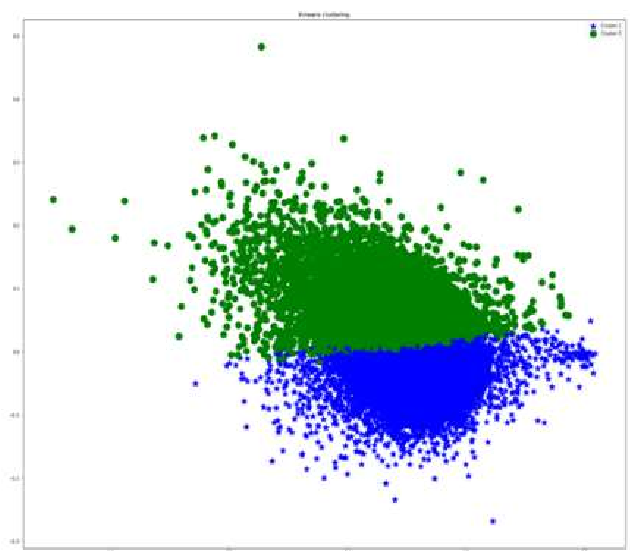

Figure 3. A Graphical Representation Of The Word2vec Vectors In A 2D Space, Where Green Labels Are Depressive Posts, And Blue Labels Are Regular Posts.

5.2.3. Gradient boosting with tf-idf

The result of applying Gradient Boosting with tf-idf with word2vec algorithm as following:

Accuracy 0.954125068904269

Metrics	precision	recall	f1-score	support
0.0	0.99	0.95	0.97	11850
1.0	0.88	0.96	0.92	4477
avg / total	0.96	0.95	0.95	16327

Precision: 96%
Recall: 95 %

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

F1-score: 95%

5.2.4. Random forest with tf-idf

The result of applying Random forest with tf-idf algorithm as following:

Accuracy 0.9553500336865315

Metrics	precision	recall	f1-score	
0	0.98	0.96	0.97	116
1	0.90	0.95	0.92	46
avg / total	0.96	0.96	0.96	163

Precision: 96%
Recall: 96 %
F1-score: 96%

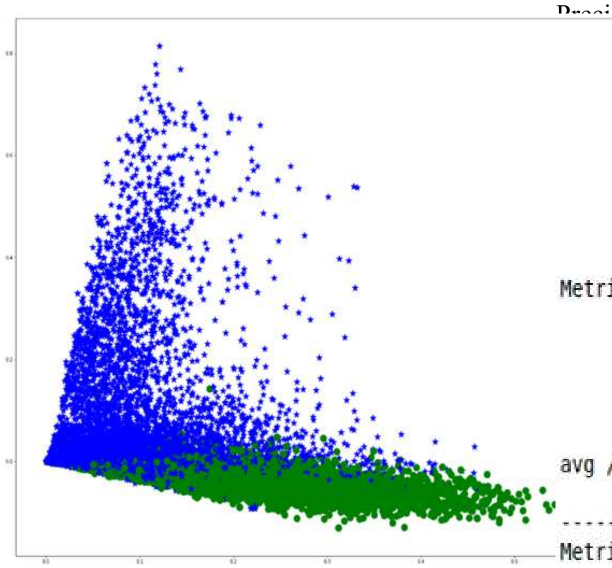

Figure 4. Graphic Representation Of Tf-Idf Vectors In 2D-Space, Where Green Marks Are Depressive Posts, And Blue Marks Are Normal Posts

5.3. Unsupervised Learning Algorithms

For the experiment, K-means with tf-idf and word2vec were tested:
1. K-means with word2vec
2. K-means with tf-idf

5.3.1. K-means with word2vec

The result of applying K-means c word2vec as following:

Metrics on training set

	precision	recall	f1-score	support
0	0.73	0.85	0.78	39268
1	0.69	0.52	0.60	26037
avg / total	0.71	0.72	0.71	65305

Metrics on testing set

	precision	recall	f1-score	support
0	0.73	0.85	0.79	9801
1	0.71	0.53	0.60	6526
avg / total	0.72	0.72	0.71	16327

ion: 72%
: 72%
re: 71%

K-means c tf-idf

3 shows the results that were taken by ng unsupervised learning algorithms as k- with tf-idf.

Metrics on training set

	precision	recall	f1-score	support
0	0.95	0.91	0.93	47763
1	0.79	0.88	0.83	17542
avg / total	0.91	0.90	0.91	65305

Metrics on testing set

	precision	recall	f1-score	support
0	0.96	0.91	0.93	11975
1	0.79	0.88	0.83	4352
avg / total	0.91	0.91	0.91	16327

Precision: 91%
Recall: 91%
F1-score: 91%

In addition, the main characteristics of results as accuracy, precision, recall, and F1-score were given in Table 1.

Journal of Theoretical and Applied Information Technology
30th November 2019. Vol.97. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Supervised Learning Algorithm For Suicidal Ideation Detection

Algorithm	Accuracy	Precision	Recall	F1-score
Gradient Boosting word2vec	90%	91%	91%	91%
Random Forest with word2vec	89	91	90	90
Gradient Boosting with tf-idf	95	96	95	95
Random Forest with tf-idf	96	96	96	96

Table 1 confirm that Gradient Boosting with tf-idf and Random Forest with tf-idf are the best classifiers for the given problem. The best supervised learning algorithm for suicidal ideation detection is Random Forest with tf-idf with 96% accuracy.
Comparison of the results of different algorithms f1-score, we can see that the Random Forest with tf-idf algorithm shows a result of 95%, which is a very good result for the given task.

To make sure that our algorithm is correct, the Receiver Operating Characteristic (ROC[1]) curve with cross-validation was built. ROC curve was applied to understand a performance measurement for classification problem at various thresholds settings.
The "steepness" of ROC curves is also important, since it is ideal to maximize the true positive rate while minimizing the false positive rate.
Figure 5 shows the ROC curve of different train and test datasets, created from K-fold cross-validation. Taking all of these curves, it is possible to calculate the mean area under curve, and see the variance of the curve when the training set is split into different subsets. This roughly shows how the classifier output is affected by changes in the training data, and how different the splits generated by K-fold cross-validation are from one another.
According to the graph, we see a stable result and we can be sure that the algorithm is well-trained to identify depressive posts.

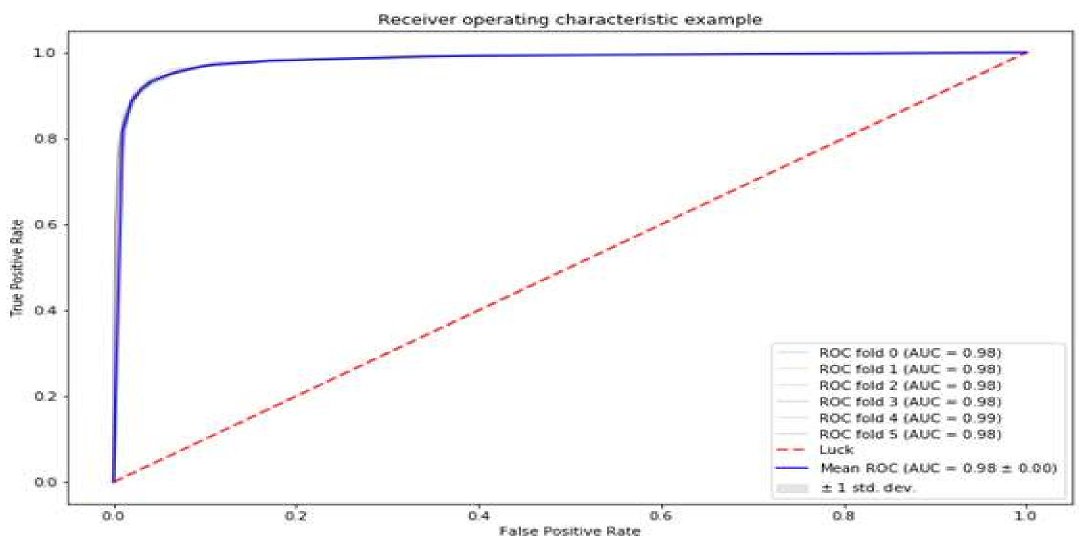

Figure 5. Roc[i] Curves With Cross-Validation

6. CONCLUSIONS

In this paper, we implemented different algorithms of supervised and unsupervised learning methods. We obtained f1-score 95% and ROC-area 0.98 with Random Forest with tf-idf vectorization model. By comparing with our previously built algorithm we increased prediction by almost 20%. We also tested how unsupervised model will perform on that dataset and it surprisingly showed great results.
There are several interesting directions of future work. One of them is to implement deep learning models with PyTorch framework. An alerting system will be built for the government to monitor emotional state of a person to prevent possible suicide attempts or any self-inflicting injuries.
We raised a very foundational research question

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

about determining the depressive posts in social media and concerned about anonymity of the data, especially when the topic is sensitive and ambiguous. We controlled parameters of training algorithms, validated it with ROC curve, and visualized results in 2D space. In addition, we made it open-source project, for future commits and changes.

In this regard we have achieved our initial goal. In the next phase of our research, we are going to apply audio, video and text analysis to identify depressed and suicidal people on the social network. We are also going to publish collected data from social media with suicidal, depressive and neutral messages that contain suicidal keywords in a data paper as a data for machine learning purpose to identify suicidal and depressive posts in social networks.

6. ACKNOWLEDGEMENTS

This research was supported by grant of the program of Ministry of Education of the Republic of Kazakhstan BR05236699 "Development of a digital adaptive educational environment using Big Data analytics". We thank our colleagues from Suleyman Demirel University (Kazakhstan) who provided insight and expertise that greatly assisted the research. We express our hopes that they will agree with the conclusions and findings of this paper.

REFERENCES:

[1] World Health Organization. (2014). Preventing suicide: A global imperative. Geneva, Switzerland. Retrieved from http://apps.who.int/iris/bitstream/10665/131056/1/9789241564779_eng.pdf?ua=1&ua=1

[2] VK.com – Vkontakte Social Network

[3] https://vk.com/topic-78863260_30603285 - About VK Social Network

[4] Brown, J. (2018). Is social media bad for you? The evidence and the unknowns. BBC Future, Jan.

[5] Naslund, J. A., Aschbrenner, K. A., & Bartels, S. J. (2016). How people with serious mental illness use smartphones, mobile apps, and social media. Psychiatric rehabilitation journal, 39(4), 364.

[6] Escobar-Viera, C. G., Shensa, A., Bowman, N. D., Sidani, J. E., Knight, J., James, A. E., & Primack, B. A. (2018). Passive and active social media use and depressive symptoms among United States adults. Cyberpsychology, Behavior, and Social Networking, 21(7), 437-443.

[7] Waytz, A., & Gray, K. (2018). Does online technology make us more or less sociable? A preliminary review and call for research. Perspectives on Psychological Science, 13(4), 473-491.

[8] Villalonga-Olives, E., & Kawachi, I. (2017). The dark side of social capital: A systematic review of the negative health effects of social capital. Social science & medicine, 194, 105-127.

[9] Wang, J. L., Gaskin, J., Rost, D. H., & Gentile, D. A. (2018). The reciprocal relationship between passive social networking site (SNS) usage and users' subjective well-being. Social Science Computer Review, 36(5), 511-522.

[10] Elhai, J. D., Tiamiyu, M., & Weeks, J. (2018). Depression and social anxiety in relation to problematic smartphone use: The prominent role of rumination. Internet Research, 28(2), 315-332.

[11] Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. Social Issues and Policy Review, 11(1), 274-302.

[12] Woods, H. C., & Scott, H. (2016). # Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. Journal of adolescence, 51, 41-49.

[13] Primack, B. A., Shensa, A., Escobar-Viera, C. G., Barrett, E. L., Sidani, J. E., Colditz, J. B., & James, A. E. (2017). Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among US young adults. Computers in human behavior, 69, 1-9.

[14] Frison, E., & Eggermont, S. (2016). Exploring the relationships between different types of Facebook use, perceived online social support, and adolescents' depressed mood. Social Science Computer Review, 34(2), 153-171.

[15] Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. Clinical Psychological Science, 6(1), 3-17.

[16] Andreassen, C. S., Billieux, J., Griffiths, M. D., Kuss, D. J., Demetrovics, Z., Mazzoni, E., & Pallesen, S. (2016). The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: A large-scale cross-sectional study. Psychology of Addictive Behaviors, 30(2), 252.

[17] Tan, Y., Chen, Y., Lu, Y., & Li, L. (2016). Exploring associations between problematic internet use, depressive symptoms and sleep disturbance among southern Chinese adolescents. International journal of environmental research and public health, 13(3), 313.

[18] Primack, B. A., Shensa, A., Escobar-Viera, C. G., Barrett, E. L., Sidani, J. E., Colditz, J. B., & James, A. E. (2017). Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among US young adults. Computers in human behavior, 69, 1-9.

[19] Lin, L. Y., Sidani, J. E., Shensa, A., Radovic, A., Miller, E., Colditz, J. B., ... & Primack, B. A. (2016). Association between social media use and depression among US young adults. Depression and anxiety, 33(4), 323-331.

[20] Hswen, Y., Naslund, J. A., Brownstein, J. S., & Hawkins, J. B. (2018). Online communication about depression and anxiety among twitter users with schizophrenia: preliminary findings to inform a digital phenotype using social media. Psychiatric Quarterly, 89(3), 569-580.

[21] Fox, J., & Moreland, J. J. (2015). The dark side of social networking sites: An exploration of the relational and psychological stressors associated with Facebook use and affordances. Computers in human behavior, 45, 168-176.

[22] Ersoy, M. (2019). Social media and children. In Handbook of Research on Children's Consumption of Digital Media (pp. 11-23). IGI Global.

[23] Olenik-Shemesh, D., & Heiman, T. (2017). Cyberbullying victimization in adolescents as related to body esteem, social support, and social self-efficacy. The Journal of genetic psychology, 178(1), 28-43.

[24] Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. Current Opinion in Behavioral Sciences, 18, 43-49.

[25] Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., ... & Zhu, W. (2017, August). Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution. In IJCAI (pp. 3838-3844).

[26] Yazdavar, A. H., Al-Olimat, H. S., Ebrahimi, M., Bajaj, G., Banerjee, T., Thirunarayan, K., ... & Sheth, A. (2017, July). Semi-supervised approach to monitoring clinical depressive symptoms in social media. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (pp. 1191-1198). ACM.

[27] Omarov, B., Omarov, N., Sarbasova, A., Sapakova, S., Koishiyeva, T., Kaliyev, N., Koishibayev, A., Omarov, B. Face recognition in video surveillance systems as an application of smart cities. Journal of Theoretical and Applied Information Technology. Volume 97, Issue 9, 15 May 2019, Pages 2572-2583.

[28] Islam, M. R., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression detection from social network data using machine learning techniques. Health information science and systems, 6(1), 8.

[29] Tiggemann, M., Hayden, S., Brown, Z., & Veldhuis, J. (2018). The effect of Instagram "likes" on women's social comparison and body dissatisfaction. Body image, 26, 90-97.

[30] De Vries, D. A., & Vossen, H. G. (2019). Social Media and Body Dissatisfaction: Investigating the Attenuating Role of Positive Parent–Adolescent Relationships. Journal of youth and adolescence, 48(3), 527-536.

[31] Tiggemann, M., & Barbato, I. (2018). "You look great!": The effect of viewing appearance-related Instagram comments on women's body image. Body image, 27, 61-66.

[32] Glodstein, S. L., DiMarco, M., Painter, S., & Ramos-Marcuse, F. (2018). Advanced practice registered nurses attitudes toward suicide in the 15-to 24-year-old population. Perspectives in psychiatric care, 54(4), 557-563.

[33] Ahad, R., & Shah, S. A. (2018). Prevalence of Suicidal Ideation and Attempts among youth of Srinagar district of J&K. AGU International Journal Of Research In Social

Science & Humanities.

[34] Omarov, B. (2017, October). Applying of audioanalytics for determining contingencies. In 2017 17th International Conference on Control, Automation and Systems (ICCAS) (pp. 744-748). IEEE.

[35] Kapka-Skrzypczak, L. (2019). Prevalence and selected risk factors of suicidal ideation, suicidal tendencies and suicide attempts in young people aged 13–19 years. Annals of Agricultural and Environmental Medicine, 26(2), 329-336.

[36] Onalbek, Z. K., Berkimbayev, K. M., Mukhamedzhanov, B. K., Omarov, B. S., Usenbek, R. R., Kendzhaeva, B. B., & Mukhamedzhanova, M. Z. (2013). Forming of Professional Competence of Future Tyeacher-Trainers as a Factor of Increasing the Quality. Middle-East Journal of Scientific Research, 15(9), 1272-1276.

[37] Memon, A. M., Sharma, S. G., Mohite, S. S., & Jain, S. (2018). The role of online social networking on deliberate self-harm and suicidality in adolescents: A systematized review of literature. Indian journal of psychiatry, 60(4), 384.

[38] Huang, X., Zhang, L., Chiu, D., Liu, T., Li, X., & Zhu, T. (2014, December). Detecting suicidal ideation in Chinese microblogs with psychological lexicons. In 2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (pp. 844-849). IEEE.

[39] Goh, T. T., & Huang, Y. P. (2009). Monitoring youth depression risk in Web 2.0. Vine, 39(3), 192-202.

[40] Månsson, K. N., Frick, A., Boraxbekk, C. J., Marquand, A. F., Williams, S. C. R., Carlbring, P., ... & Furmark, T. (2015). Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Translational psychiatry*, *5*(3), e530.

[41] Just, M. A., Pan, L., Cherkassky, V. L., McMakin, D. L., Cha, C., Nock, M. K., & Brent, D. (2017). Nature human behaviour, 1(12), 911.

[42] Huang, X., Zhang, L., Chiu, D., Liu, T., Li, X., & Zhu, T. (2014, December). Detecting suicidal ideation in Chinese microblogs with psychological lexicons. In 2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (pp. 844-849). IEEE.

[43] Van der Feltz-Cornelis, C. M., Sarchiapone, M., Postuvan, V., Volker, D., Roskar, S., Grum, A. T., ... & Ibelshäuser, A. (2011). Best practice elements of multilevel suicide prevention strategies. Crisis.

[44] Schulberg, H. C., Bruce, M. L., Lee, P. W., Williams Jr, J. W., & Dietrich, A. J. (2004). Preventing suicide in primary care patients: the primary care physician's role. General hospital psychiatry, 26(5), 337-345.

[45] Bauer, G. R., Scheim, A. I., Pyne, J., Travers, R., & Hammond, R. (2015). Intervenable factors associated with suicide risk in transgender persons: a respondent driven sampling study in Ontario, Canada. BMC public health, 15(1), 525.

[46] Fonseka, T. M., Bhat, V., & Kennedy, S. H. (2019). The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. Australian & New Zealand Journal of Psychiatry, 0004867419864428.

[47] Aladağ, A. E., Muderrisoglu, S., Akbas, N. B., Zahmacioglu, O., & Bingol, H. O. (2018). Detecting suicidal ideation on forums: proof-of-concept study. Journal of medical Internet research, 20(6), e215.

[48] Berrouiguet, S., Barrigón, M. L., Castroman, J. L., Courtet, P., Artés-Rodríguez, A., & Baca-García, E. (2019). Combining mobile-health (mHealth) and artificial intelligence (AI) methods to avoid suicide attempts: the Smartcrises study protocol. BMC psychiatry, 19(1), 1-9.

[49] de Andrade, N. N. G., Pawson, D., Muriello, D., Donahue, L., & Guadagno, J. (2018). Ethics and artificial intelligence: suicide prevention on facebook. Philosophy & Technology, 31(4), 669-684.

[50] O'Keefe, V. M., Tucker, R. P., Cole, A. B., Hollingsworth, D. W., & Wingate, L. R. (2018). Understanding indigenous suicide through a theoretical lens: A review of general, culturally-based, and indigenous frameworks. Transcultural psychiatry, 55(6), 775-799.